CLAIMS

What is claimed is:

1. A nucleic acid reference library comprising a heterogeneous mixture of nucleic acid fragments, wherein said fragments comprises a portion of a polymorphic subregion of a polymorphic consensus sequence, wherein each of said polymorphic subregions is bounded by first restriction sites and comprises an internal polymorphic restriction site which is different from said first site and wherein said library is enriched for fragments other than those located between said polymorphic subregions.

2. A nucleic acid reference library according to claim 1 wherein at least a subpopulation of said library comprises nucleic acid fragments further comprising oligonucleotide tags wherein different nucleic acid fragments are linked to different oligonucleotide tags.

3. The nucleic acid reference library according to claim 2 further comprising a replicable vector.

4. The nucleic acid reference library according to claim 2 wherein said oligonucleotide tags comprise oligonucleotides of the form:

$$S_1S_2S_3 \ldots S_n$$

wherein each of S_1 through S_n are subunits consisting of an oligonucleotide having a length from 3 to 9 nucleotides and are selected from a minimally cross-hybridizing set, n is in the range of from 4 to 10, and wherein said tag has a length in the range of from 12 to 60 nucleotides or base pairs.

5. A composition comprising subpopulations of microparticles wherein each subpopulation comprises at least one microparticle comprising a polymorphic probe wherein the polymorphic probes of the different subpopulations are different from said others and comprise a portion of a polymorphic subregion of a polymorphic consensus sequence.

6. The composition of claim 5 wherein each of said subpopulations further comprises a unique oligonucleotide tags.

7. The composition according to claim 6 wherein said oligonucleotide tags are positioned between said microparticle and said polymorphic probe.

8. The composition according to claim 6 wherein said oligonucleotide tags comprise oligonucleotides of the form:

$$S_1S_2S_3 \ldots S_n$$

wherein each of S_1 through S_n are subunits consisting of an oligonucleotide having a length from 3 to 9 nucleotides and are selected from a minimally cross-hybridizing set, n is in the range of from 4 to 10, and wherein said tag has a length in the range of from 12 to 60 nucleotides or base pairs.

9. An array comprising a solid support having defined regions on the surface thereof, wherein each region comprises a different polymorphic probe, and wherein each of said polymorphic probes comprises a portion of a polymorphic subregion of a polymorphic consensus sequence.

10. The array of claim 9 wherein each of said regions further comprises an oligonucleotide tag.

11. The array of claim 11 wherein said oligonucleotide tags are positioned between said surface and said polymorphic probe.

12. The array according to claim 11 wherein said oligonucleotide tag comprises oligonucleotides of the form:

$$S_1S_2S_3 \ldots S_n$$

wherein each of S_1 through S_n are subunits consisting of an oligonucleotide having a length from 3 to 9 nucleotides and are selected from a minimally

cross-hybridizing set, n is in the range of from 4 to 10, and wherein said tag has a length in the range of from 12 to 60 nucleotides or base pairs.

13. A nucleic acid reference library of genomic DNA from a plurality of individuals comprising:

a heterogeneous mixture of restriction fragments of a first restriction endonuclease;

wherein said restriction fragments from the same locus contain members having one or more restriction site polymorphisms with respect to a second restriction endonuclease, the number of such members forming a proper subset of the total number of restriction fragments from said locus.

14. A nucleic acid reference library of genomic DNA from a plurality of individuals comprising:

a heterogeneous mixture of restriction fragments of a first restriction endonuclease;

wherein said restriction fragments from the same locus contain at least one member having one or more restriction site polymorphisms with respect to a second restriction endonuclease and at least one member without said restriction site polymorphism.

15. A method of making a reference library comprising a mixture of heterogeneous nucleic acid fragments comprising:

digesting pooled nucleic acid comprising first restriction sites with a first restriction endonuclease to produce a mixture of restriction fragments;

forming a first population of single stranded nucleic DNA fragments from a first subpopulation of said restriction fragments, wherein said first subpopulation comprises a portion of said mixture of restriction fragments which comprises a second restriction site which is different from said first restriction site; and

forming a second population of single stranded DNA fragments from a second subpopulation of said restriction fragments wherein said second subpopulation comprises a portion of said restriction fragments that do not contain said second restriction site, and wherein said first single stranded DNA fragments from said first subpopulation have complementary sequences to said second single stranded DNA fragments from said second subpopulation whenever said single stranded DNA fragments are derived from the same restriction fragment;

hybridizing the first and second populations of single stranded DNA fragments to form a population of duplexes; and

isolating said duplexes to form a reference population of restriction fragments.

16. The method of Claim 16 further comprising the step of pretreating said pooled nucleic acid to enrich for non-repetitive sequences.

17. A method for determining the ratio of a polymorphic subregion between at least two different pools of test nucleic acid comprising

generating a first pool of restriction endonuclease fragments from a first pool of test nucleic acids comprising first restriction sites by digesting said pool with a first restriction endonuclease;

generating a second pool of restriction endonuclease fragments from a second pool of test nucleic acids comprising first restriction sites by digesting said pool with said first restriction endonuclease;

enriching said first and said second pools of restriction fragments for those fragments which contain a polymorphism associated with a second restriction site to form first and second enriched populations;

contacting said first and said second enriched populations with a reference library comprising probes enriched for subregions which are polymorphic for said second restriction site; and

determining the ratio of binding of said probes with said first and said second enriched populations.

18. The method of claim 18 wherein said first pool of test nucleic acids is from a population of individuals having a first phenotype and said second pool of test nucleic acids is from a population of individuals having a second phenotype.